Sex determination of human remains from peptides in tooth enamel

Nicolas Andre Stewart[1], Raquel Gerlach[2], Rebecca Gowland[3], Kurt Gron[3], Janet Montgomery[3]

[1]University of Brighton, [2]University of Sao Paulo, [3]Durham University

Submitted to Proceedings of the National Academy of Sciences of the United States of America

The assignment of biological sex to archaeological human skeletons is a fundamental requirement for the reconstruction of the human past. It is conventionally and routinely performed on adults using metric analysis and morphological traits arising from post-pubertal sexual dimorphism. A maximum accuracy of c. 95% is possible if both the cranium and *os coxae* are present and intact, but this is seldom achievable for all skeletons. Furthermore, for infants and juveniles, there are no reliable morphological methods for sex determination without resorting to DNA analysis which requires good DNA survival and is time consuming. Consequently, sex determination of juvenile remains is rarely undertaken and a dependable and expedient method that can correctly assign biological sex to human remains of any age is highly desirable. Here, we present a method for sex determination of human remains by means of a minimally destructive surface acid-etch of tooth enamel and subsequent identification of sex chromosome-linked isoforms of amelogenin – an enamel-forming protein – by nanoLC-MS. Tooth enamel is the hardest tissue in the human body and survives burial exceptionally well, even when the rest of the skeleton or DNA in the organic fraction has decayed. Our method can reliably determine the biological sex of humans of any age using a body tissue that is difficult to cross-contaminate and is most likely to survive. The application of this method will render sex determination of adults and for the first time, juveniles, a reliable and routine activity in future bioarchaeological and medico-legal science contexts.

sex determination | tooth enamel | amelogenin | human remains | mass spectromtry

Introduction

Sex is a fundamental primary characteristic for the analysis of human skeletal remains from archaeological and medico-legal contexts. Techniques for determining other key identifying features, such as age-at-death and stature, are also sex-dependent (1). Accurate profiles of sex are crucial for reconstructing past societies in terms of demography, identity and epidemiology, but are also essential in medico-legal contexts for identifying individuals (e.g. in mass disasters). In adults, sex can be determined with relative accuracy that is usually estimated to fall between 80-95%, depending upon factors such as skeletal preservation and the degree of sexual dimorphism within the sample (2). One of the key limiting factors of osteological analyses to date has been an inability to reliably determine the sex of individuals from skeletal features prior to approximately 18 years of age (3). Numerous methods have tried to do so, usually applying studies of the morphological and metrical characteristics of the infant and juvenile mandible, dentition, and ilium, e.g. (4, 5). However, no method has proven sufficiently reliable when tested on documented skeletal samples (6).

Ancient DNA analysis is unfortunately not the answer to this problem due to issues of preservation, contamination and expense. For example, several DNA studies attempted to determine the sex of infants from Romano-British sites with the aim of assessing whether preferential female infanticide was practiced (7, 8). In all of these studies viable results were only ever obtained from a small proportion of the overall number sampled. The destructive, costly, and inconclusive nature of DNA sex determination means that it has rarely been attempted on any scale on human remains from archaeological sites.

Herein we present a method for secure biological sex determination of human remains by means of a minimally destructive surface acid etch of tooth enamel and subsequent identification of sex chromosome-linked isoforms of amelogenin– an enamel-forming protein - by nano-liquid chromatography mass spectrometry (nanoLC-MS). Tooth enamel is the hardest tissue in the human body and survives burial exceptionally well, even when the rest of the skeleton or DNA in the organic fraction has decayed. Therefore, this method holds promise to reliably determine the biological sex of humans of any age using a body tissue that is most likely to survive intact. It is minimally destructive, inexpensive, and reliable.

Results and Discussion

The acid-etch procedure to extract peptides from tooth enamel results in a complex base peak chromatogram when analysed by nanoLC-MS/MS (Fig. 1). It has previously been shown that peptides can be identified using various acid-etch methods and that these originate from the major tooth enamel proteins: amelogenin, ameloblastin and enamelin (9-12). During enamel maturation the majority of these proteins are processed by proteases resulting in peptides of varying lengths and remain in the mature enamel (9, 11, 12).

For amelogenin, peptides from the central portion of the protein are absent but peptides from the N- and C-termini remain, and these are identified (Supporting data). The dimorphic differences between amelogenin X and Y are found in these

Significance

The ability to assign biological sex to human skeletal remains is a fundamental requirement in archaeology, palaeoanthropology and medico-legal sciences. While DNA sequencing can be used, it is expensive, time-consuming and often fails due to the poor quality of the remaining DNA. An easier, more reliable and consistently applicable method is needed. We present a method for sex determination of human remains using peptides retrieved from tooth enamel. Amelogenin is an enamel-forming protein which is encoded-for by chromosomes X and Y, with slight differences in their amino acid sequences. Peptides with these differences were identified by nanoflow liquid chromatography mass spectrometry and correctly assign sex to archaeological human remains of various chronological ages, from hundreds to thousands of years old.

Reserved for Publication Footnotes

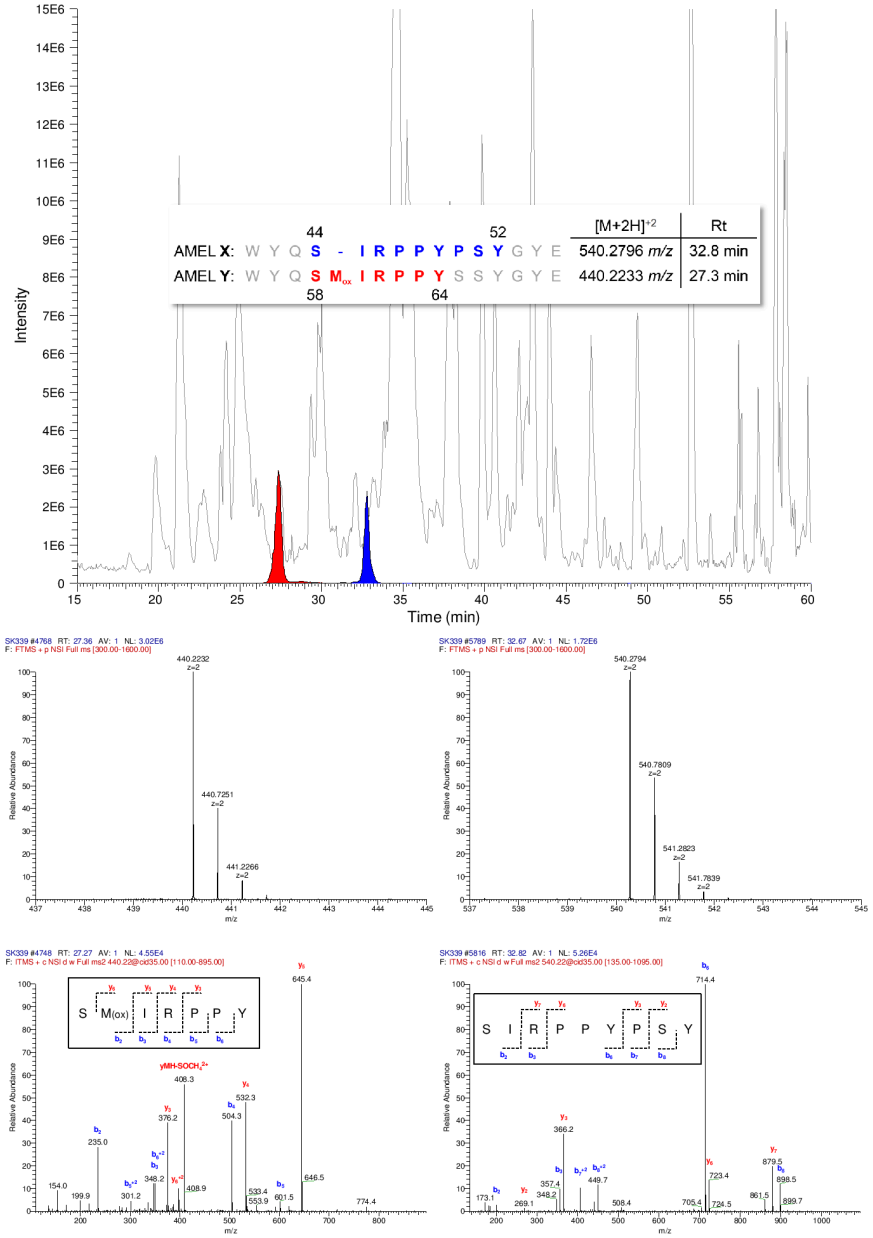

Fig. 1. A representative base peak chromatogram (300 – 1600 m/z) produced from Fewston sample SK339 with inset highlighting the amino acid sequences of the two dimorphic peptides of amelogenin; AMELY-(58-64)-peptide and AMELX-(44-52)-peptide. The reconstructed ion chromatograms (to 4 ppm) for each of these are shown in red and blue, respectively, with full scan MS and corresponding MS/MS below.

regions, and one such peptide identified is the AMELY-(58-64)-peptide which possesses an additional methionine when compared to the aligned sequence of AMELX (Fig. 1 inset). In one of the samples (SK130) the AMELX peptide seems to be relatively lower in abundance compared to the AMELY peptide. This most likely reflects a higher relative amount of the AMELY peptide, as this peptide contains a methionine, and it may be oxidized in greater amount in this sample. Therefore, the oxidized version of the AMELY peptide was chosen for "Y" sex confirmation as it is expected that this one will predominate in old samples as opposed to the unoxidized peptide which is anticipated to be low in abundance or absent.

A peptide from amelogenin X; AMELX-(44-52)-peptide, with similar ion intensity to the AMELY-(58-64)-peptide was used to clearly assign sex based on the presence or absence of AMELY-(58-64) from the seven Fewston samples (Fig. 2) and the three sets of sex paired samples (Fig. 3).

The sexual dimorphic protein isoforms of amelogenin and the fact that peptides from the dimorphic regions of these isoforms serendipitously remain in tooth enamel provides an excellent means for unequivocally establishing the sex of human remains by nanoLC-MS. The unique nature of tooth enamel, as the most dense and hardest tissue, makes it a perfect repository for these peptides. The two dimorphic peptides chosen, AMELX-(44-52)-peptide and AMELY-(58-64)-peptide, clearly differentiate between male and female, where both the AMELY and AMELX peptides are found in male samples and only the AMELX-peptide is found in female samples. In all cases our results agree with the assignment of sex by either coffin plates or standard osteological methods.

In one of the samples (SK130) the AMELX peptide seems to be relatively lower in abundance compared to the AMELY peptide. This most likely reflects a higher relative amount of the AMELY peptide, as this peptide contains a methionine, and it may be oxidized to a greater degree in this sample. It should

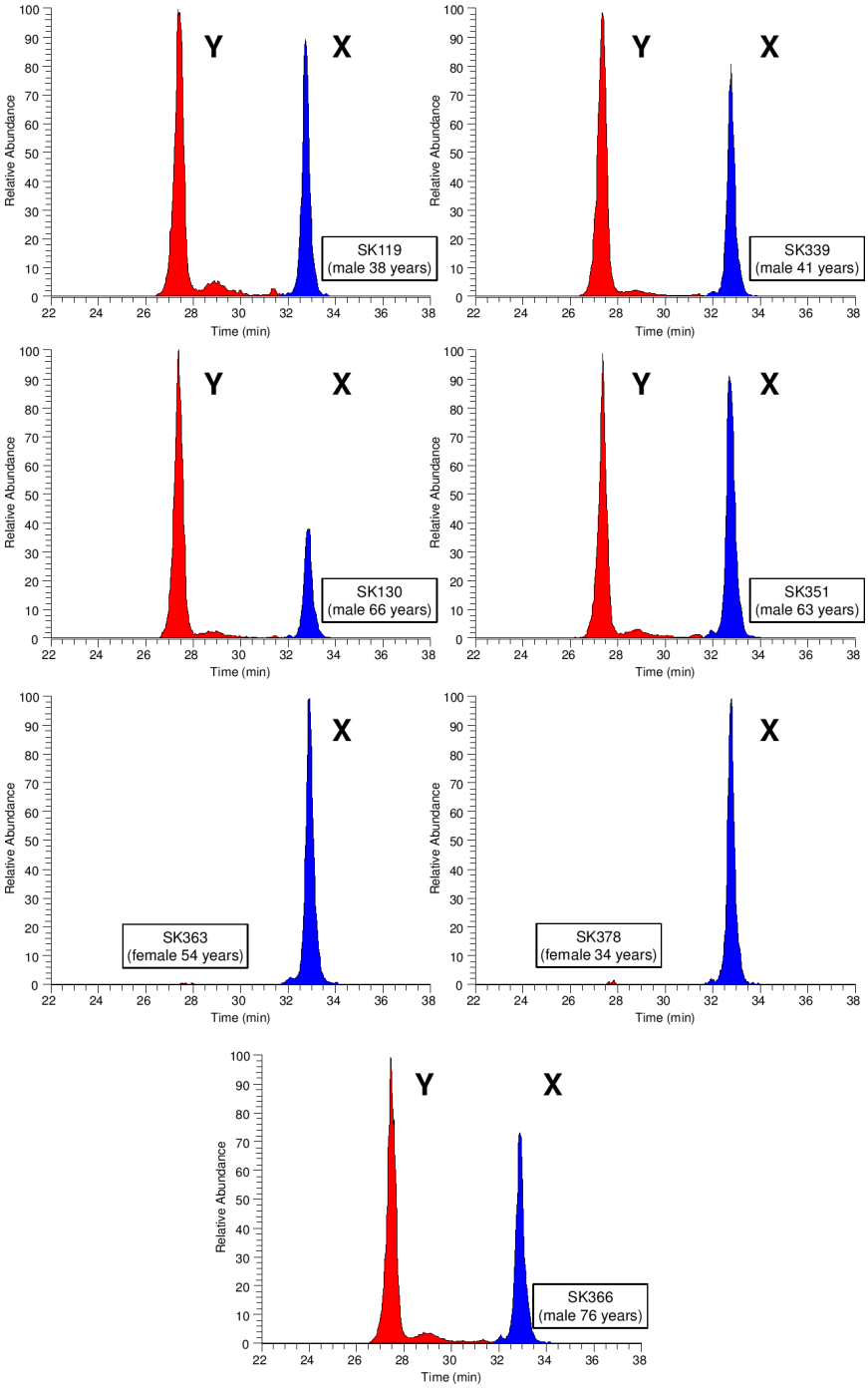

Fig. 2. Reconstructed ion chromatograms (RICs) for the AMELY-(58-64)-peptide (440.2233 m/z) and AMELX-(44-52)-peptide (540.2796 m/z) (4 ppm mass tolerance) for the seven 19th century Fewston samples. Peaks corresponding to these are shown in red and blue, respectively. Parenthetical text indicates known sex and age-at-death.

be noted that this method identifies the presence of peptides originating from sex chromosome-linked isoforms of amelogenin, and currently cannot identify polymorphisms of multiple copies of these chromosomes (e.g. aneuploid 47, XXY or 47, XYY). Quantitation of these peptides may allow for this in these rare cases and warrants further investigation.

The ability to determine the sex of infant and juvenile remains completely revolutionises studies of growth, child-care, epidemiology and demography in the past. For the first time, it will allow osteologists to examine sex-specific cultural treatment and differentiate between the health of boys and girls, as well as sex-specific growth trajectories and past developmental milestones such as age-of-puberty and subsequent repercussions for fertility. Sites with poor preservation are common in archaeological contexts and at such sites teeth generally survive better than bone and thus sex can be established for adults as well as juvenile skeletons in the absence of key skeletal identifiers. Additionally, the dimorphic peptide sequence is identical in apes (Fig S1) and so should be present in all hominins. Finally, this technique will also have a transformative effect on human identification in medico-legal contexts, such as mass disasters and war graves, allowing sex to be established both reliably and cost-effectively.

Materials and Methods

Fewston is a small village located in the Washburn Valley, near Harrogate

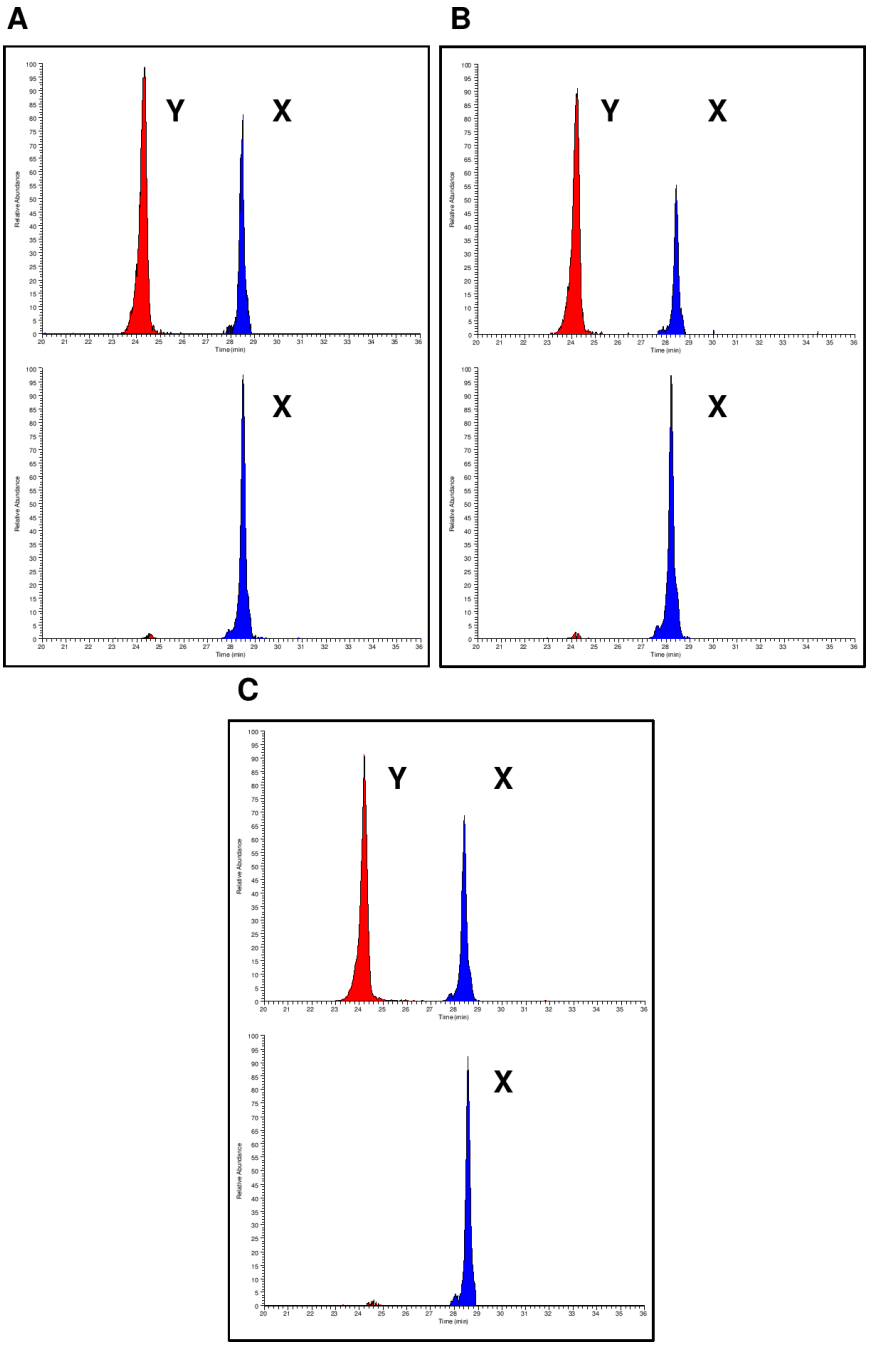

Fig. 3. Reconstructed ion chromatograms (RICs) for the AMELY-(58-64)-peptide (440.2233 m/z) and AMELX-(44-52)-peptide (540.2796 m/z) (4 ppm mass tolerance) for three male/female pairs of archaeological samples, A) from St. Guthlac's Priory 12 – 16th century AD, B) from Whitwell c. 5,700 BP and C) from Seaham 7th-9th centuries AD (previously published). Parenthetical text indicates osteological age and sex determinations.

in North Yorkshire, UK. The skeletal assemblage was excavated from the parish churchyard in advance of building work in 2009-2010 and was re-buried in September 2016. Twenty-one of the excavated individuals were confidently identified based on coffin plates and grave monuments. All of these identified individuals date to the late nineteenth century. Etches were performed on teeth from seven adult individuals of known identity and sex. Sex was assigned by coffin plates and confirmed by osteological analysis (13) (see Table 1). All samples were anonymised after removal of peptides and measured blind.

Additionally, male and female pairs from three archaeological sites ranging in date from the early Neolithic c. 5,700 BP to the medieval period were tested to determine survival and recovery of sufficient proteins over archaeological timescales and a variety of burial contexts. For each pair, sex had been previously assigned using standard osteological methods as describe in Table 1. These were not analysed blind.

Whitwell, Seaham and St Guthlac's samples were prepared and analysed as previously described (9). All reagents used were of analytical grade and solvents used for nanoLC-MS/MS were MS grade. The Fewston samples were prepared using a modified streamlined version of the protocol as follows. The tooth surface was abraded using a dental burr to remove obvious surface contaminants. The enamel was washed with 3% H_2O_2 for 30 sec, and rinsed with ultrapure water (Elga Purelab Ultra, 18.2 MΩ-cm). In the cap of a 0.2 mL Eppendorf tube ca. 60 µL of 5% (v/v) HCl was placed leaving a convex meniscus protruding above the lip. An initial etch was performed by lowering the tooth onto the HCl and retaining contact for two minutes. This first etch was discarded. A second two-minute etch was then performed and retained as the etch solution. In an adjustable 0.5-10 µL pipette set to 10 µL a C18 resin loaded ZipTip (EMD Millipore, Billerica, MA, USA, ZTC18S096) was conditioned three times with 100% acetonitrile, then three times with 0.1% (v/v) formic acid discarding each draw. The etch solution was bound to the ZipTip by pipetting the solution up and down ten times, discarding the last draw. The ZipTip was washed six times with 0.1% (v/v) formic acid, discarding each wash. The adjustable pipette was set to 4 µL and the resin-bound peptides were eluted into a 4 µL 60%

Table 1.

Site Location	Period	Type of burial	Skeleton No.	Age and Sex[1]	Methods used to determine sex	Reference
Whitwell, Derbyshire, UK	Neolithic c. 5,700 BP	Fragmentary, disarticulated, cranium (SK485) and articulating mandible (SK219)	SK219	Adult Female	Morphological traits of the mandible of SK219 and the articulating maxilla of SK485: (3, 14); Marked sexual dimorphism.	(15)
Whitwell, Derbyshire, UK	Neolithic c. 5,700 BP	Fragmentary, disarticulated mandible	SK534	Adult Male	Morphological traits of the mandible only: (3, 14); Marked sexual dimorphism.	(15)
Seaham, County Durham, UK	7^{th} – 9^{th} centuries AD	Inhumation cemetery	FFS SK15	Female 26-45 years	Morphological traits of the pelvis and skull: (3, 14, 16)	(17)
Seaham, County Durham, UK	7^{th} – 9^{th} centuries AD	Inhumation cemetery	FFS SK3	Male 36+ years	Morphological traits of the pelvis and skull: (3, 14, 16)	(17)
St Guthlac's Priory, Hereford, UK	12^{th} – 16^{th} century AD	Inhumation cemetery	SK 9503	Female Old Adult	Morphological traits of the pelvis and skull: (3, 14, 16)	(18)
St Guthlac's Priory, Hereford, UK	12^{th} – 16^{th} century AD	Inhumation cemetery	SK 9515	Male Old Adult	Morphological traits of the pelvis and skull: (3, 14, 16) Significant grave goods: chalice	(18)
Fewston, North Yorkshire, UK	19^{th} century AD	Inhumation cemetery	SK363	Female 54 years	Documented age and sex: coffin plate. Morphological traits of the pelvis and skull: (2, 3, 16)	(13)
Fewston, North Yorkshire, UK	19^{th} century AD	Inhumation cemetery	SK378	Female 34 years	Documented age and sex: coffin plate. Morphological traits of the pelvis and skull: (2, 3, 16)	(13)
Fewston, North Yorkshire, UK	19^{th} century AD	Inhumation cemetery	SK366	Male 76 years	Documented age and sex: coffin plate. Morphological traits of the pelvis and skull: (2, 3, 16)	(13)
Fewston, North Yorkshire, UK	19^{th} century AD	Inhumation cemetery	SK119	Male 38 years	Documented age and sex: coffin plate. Morphological traits of the pelvis and skull: (2, 3, 16)	(13)
Fewston, North Yorkshire, UK	19^{th} century AD	Inhumation cemetery	SK339	Male 41 years	Documented age and sex: coffin plate. Morphological traits of the pelvis and skull: (2, 3, 16)	(13)
Fewston, North Yorkshire, UK	19^{th} century AD	Inhumation cemetery	SK130	Male 66 years	Documented age and sex: coffin plate. Morphological traits of the pelvis and cranium: (2, 3, 16)	(13)
Fewston, North Yorkshire, UK	19^{th} century AD	Inhumation cemetery	SK351	Male 63 years	Documented age and sex: grave stone. Morphological traits of the pelvis and skull: (2, 3, 16)	(13)

[1] As previously determined by osteological, epigraphic and grave goods

Details of samples used in this study.

acetonitrile/0.1% formic acid elution buffer and lyophilized. Samples were dissolved in 12 µL 0.1% trifluoroacetic acid (TFA) in water, centrifuged on a bench top centrifuge for 5 min to remove any particulate matter and transferred (10 µL) to glass autosampler vials. Five µL was injected for analysis by reversed-phase nanoflow liquid chromatography (nanoRS U3000, Thermo Scientific) coupled to a hybrid linear ion trap Orbitrap (Orbitrap XL, Thermo Scientific) mass spectrometer. Peptides were first loaded onto a C18 trapping cartridge (Pepmap100 C18; 0.3 × 5 mm ID; 5 µm particle size) for 10 min at a flow-rate of 5 µL/min with 0.1% TFA. Separation was achieved at a flow rate of 300 nL/min on an analytical column (25 cm x 75 µm; 5 µm particle size, C18 PepMap100, Thermo Scientific) with a gradient starting at 1% and increased to 13.3% solvent B over 20 min, then to 25.6% over 15 min, 45% over 10 min, 99% over 15 min, held constant at 99% for 5 min, returned to 1% and equilibrated for 20 min. Nanoelectrospray ionisation (nano-ESI) was performed with a 10 µm uncoated silica tip emitter (New Objective, FS360-20-10-N-20). The MS was operated in data-dependent MS/MS mode in which each full MS scan was collected in the Orbitrap (300–1600 *m/z*, R = 60,000 @ 400 *m/z*) followed with up to nine MS/MS scan events performed in the linear ion trap where the most abundant peptide molecular ions were selected for collision-induced dissociation (CID), using a normalized collision energy of 35%. Total MS acquisition time was 64 min. Data was searched against the human proteome (UniprotKB, 10/15) using MaxQuant (V 1.5.1.2) employing default search settings with methionine oxidation as variable modification, unspecific digestion mode and minimum peptide length of 6. The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE (19) partner repository with the dataset identifier PXD007856.

Acknowledgements

We thank Anwen Caffell, Malin Holst (York Osteoarchaeology Ltd), John Buglass, The Heritage Lottery Fund and the Washburn Heritage Centre for access to, and information regarding, the Fewston skeletal assemblage. We also thank Andrew Chamberlain and Charlotte Roberts for access to the Whitwell and St Guthlac's skeletal collections and for their unpublished sex determinations of the individuals included herein. We are also grateful to Niall Hammond (Archaeo-Environment Ltd) and Northern Archaeological Associates for access to the Seaham skeletal collection excavated in 1999 and to Tessi Loeffelmann for sex determination of these individuals in the course of her MSc dissertation in 2014 'Of doomed and transient men: the Anglo-Saxon skeletons of Seaham in their context' at Durham University. KG thanks the Leverhulme Trust for funding. RFG acknowledges support from the State of Sao Paulo Research Foundation (FAPESP) and the Brazilian Research Council (CNPq) for studies on enamel peptides and trace elements. Lastly, we thank Dr Steve M. Sweet at the University of Sussex for access to the nanoLC-Orbitrap MS instrument.

1. Gowland R & Thompson TJU (2013) The skeleton. *Human identity and identification*, (Cambridge University Press, Cambridge ; New York), p 225.
2. Cox M & Mays S (2000) *Human osteology in archaeology and forensic science* (Greenwich Medical Media, London) pp 117-120.
3. Haas J, Buikstra JE, Ubelaker DH, Aftandilian D, & Field Museum of Natural History. (1994) *Standards for data collection from human skeletal remains : proceedings of a seminar at the Field Museum of Natural History, organized by Jonathan Haas* (Arkansas Archeological Survey, Fayetteville, AR).
4. Schutkowski H (1993) Sex determination of infant and juvenile skeletons: I. Morphognostic features. *Am J Phys Anthropol* 90(2):199-205.
5. Vlak D, Roksandic M, & Schillaci MA (2008) Greater sciatic notch as a sex indicator in juveniles. *Am J Phys Anthropol* 137(3):309-315.
6. Lewis M (2007) *The bioarchaeology of children : perspectives from biological and forensic anthropology/ Mary E. Lewis* (Cambridge University Press, Cambridge).
7. Abu-Mandil Hassan N, Brown KA, Eyers J, Brown TA, & Mays S (2014) Ancient DNA study of the remains of putative infanticide victims from the Yewden Roman villa site at Hambleden, England. *Journal of Archaeological Science* 43:192-197.
8. Mays S & Faerman M (2001) Sex Identification in Some Putative Infanticide Victims from Roman Britain Using Ancient DNA. *Journal of Archaeological Science* 28(5):555-559.
9. Stewart NA, *et al.* (2016) The identification of peptides by nanoLC-MS/MS from human surface tooth enamel following a simple acid etch extraction. *RSC Advances* 6(66):61673-61679.
10. Castiblanco GA, *et al.* (2015) Identification of proteins from human permanent erupted enamel. *Eur J Oral Sci* 123(6):390-395.
11. Porto IM, Laure HJ, de Sousa FB, Rosa JC, & Gerlach RF (2011) New techniques for the recovery of small amounts of mature enamel proteins. *Journal of Archaeological Science* 38(12):3596-3604.
12. Porto IM, *et al.* (2011) Recovery and identification of mature enamel proteins in ancient teeth. *Eur J Oral Sci* 119 Suppl 1:83-87.
13. Caffell A & Holst M (2010) Osteological Analysis, The Church of Saint Michael and Saint Lawrence, Fewston. in *Library of Unpublished Fieldwork Reports* (https://doi.org/10.5284/1025701).
14. Bass WM (1987) *Human osteology : a laboratory and field manual* (Missouri Archaeological Society, Columbia, Mo.) 3rd Ed pp xviii, 327 p.
15. Chamberlain AT & Witkin AV (2011) A Neolithic cairn at Whitwell, Derbyshire. *Derbyshire archaeological journal* 131:1-131.
16. Brickley M & McKinley JI (2004) *Guidelines to the Standards for Recording Human Remains* (Institute of Field Archaeologists Paper No. 7).
17. Loeffelmann T (2014) 'Of doomed and transient men: the Anglo-Saxon skeletons of Seaham in their context. MSc (Durham).
18. Roberts MT (2005) *Shropshire Parish registers. Diocese of Hereford. Vol. XXIV Pt. 5, Bishop's Castle baptisms and burials 1791-1837*.
19. Vizcaino JA, *et al.* (2016) 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* 44(22):11033.

Figure 1. A representative base peak chromatogram (300 – 1600 m/z) produced from Fewston sample SK339 with inset highlighting the amino acid sequences of the two dimorphic peptides of amelogenin; AMELY-(58-64)-peptide and AMELX-(44-52)-peptide. The reconstructed ion chromatograms (to 4 ppm) for each of these are shown in red and blue, respectively, with full scan MS and corresponding MS/MS below.

Figure 2. Reconstructed ion chromatograms (RICs) for the AMELY-(58-64)-peptide (440.2233 m/z) and AMELX-(44-52)-peptide (540.2796 m/z) (4 ppm mass tolerance) for the seven 19th century Fewston samples. Peaks corresponding to these are shown in red and blue, respectively. Parenthetical text indicates known sex and age-at-death.

Figure 3. Reconstructed ion chromatograms (RICs) for the AMELY-(58-64)-peptide (440.2233 m/z) and AMELX-(44-52)-peptide (540.2796 m/z) (4 ppm mass tolerance) for three male/female pairs of archaeological samples, A) from St. Guthlac's Priory 12 – 16th century AD, B) from Whitwell c. 5,700 BP and C) from Seaham 7th-9th centuries AD (previously published). Parenthetical text indicates osteological age and sex determinations.